Toward A Simple Diagnostic Index for Acute Uncomplicated Urinary Tract Infections

Bart J. Knottnerus, MD, PhD[1]

Suzanne E. Geerlings, MD, PhD[2]

Eric P. Moll van Charante, MD, PhD[1]

Gerben ter Riet, MD, PhD[1]

[1]Department of General Practice, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands

[2]Department of Internal Medicine / Infectious Diseases, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands

MORE ONLINE
www.annfammed.org

ABSTRACT

PURPOSE Whereas a diagnosis of acute uncomplicated urinary tract infection (UTI) in clinical practice comprises a battery of several diagnostic tests, these tests are often studied separately (in isolation from other test results). We wanted to determine the value of history and urine tests for diagnosis of uncomplicated UTIs, taking into account their mutual dependencies and information from preceding tests.

METHODS Women with painful and/or frequent micturition answered questions about their signs and symptoms (history) of UTIs and underwent urine tests. A culture was the reference standard (10^3 colony-forming units per milliliter). A diagnostic index was derived using logistic regression with bootstrapped backward selection and parameter-wise shrinkage. Risk thresholds for UTI of 30% and 70% were used to analyze discriminative properties. Six models were compared: (1) history only, (2) history + urine dipstick, (3) history + urine dipstick + urinary sediment, (4) history + urine dipstick + dipslide, and (5) history + urine dipstick + urinary sediment + dipslide; we then added (6) a test only for patients with an intermediate risk (between 30% and 70%) after the preceding test.

RESULTS One hundred ninety-six women were included (UTI prevalence 61%). Seven variables were selected from history (3), dipstick (2), sediment (1), and dipslide (1). History correctly classified 56% of patients as having a UTI risk of either <30% or >70%. History and urine dipstick raised this to 73%. The 3 models with the addition of urinary sediment and dipslide, separately and in combination, performed hardly better. The sixth model, in which those at intermediate risk after history and received an additional test, correctly classified 83%. The patient's suspicion of a UTI and a positive nitrite test were the strongest indicators of a UTI.

CONCLUSIONS Most women with painful and/or frequent micturition can be correctly classified as having either a low or a high risk of UTI by asking 3 questions: Does the patient think she has a UTI? Is there at least considerable pain on micturition? Is there vaginal irritation? Other women require additional urine dipstick investigation. Sediment and dipslide have little added value. External validation of these recommendations is required before they are implemented in practice.

AJC *Annals* Journal Club selection; see inside back cover or http://www.annfammed.org/AJC/.

CORRESPONDING AUTHOR

Bart J. Knottnerus, MD, PhD
Department of General Practice
Academic Medical Center
University of Amsterdam
PO Box 22700, 1100 DE
Amsterdam, The Netherlands
b.j.knottnerus@amc.uva.nl

INTRODUCTION

Acute uncomplicated urinary tract infections (UTIs) are infections of the lower urinary tract in healthy, nonpregnant, adult women. The diagnosis is made by the presence of urinary symptoms in combination with bacteriuria.[1-4] Sixty percent of all women experience at least 1 UTI during their life.[5] The symptoms are bothersome and have a negative impact on quality of life.[6-9] Although empiric antibiotic treatment of all women with urinary symptoms has been reported to be cost-effective,[10,11] bacterial resistance is rising,[12-15] and an accurate diagnosis of UTI is needed to facilitate a well-targeted use of antibiotics.

Various medical history questions and urine investigations can be used for UTI diagnosis, of which nitrite, blood, and leukocyte esterase urine

dipstick tests, microscopic examination of the urinary sediment, and dipslide are the ones most widely applied. Most of these diagnostic indicators have been studied in single-test evaluations, implying that a test is compared with the urine culture without taking into account the results of preceding tests, including clinical history.[16-19] In clinical practice, however, the diagnostic work-up is multivariable, and test results are mutually dependent.[20-24] For example, the diagnostic value of the dipslide used in isolation may be considerable, but it might not add much once the clinical history and nitrite test result are known. As a result, performance of expensive and time-consuming tests, such as urinary sediment and dipslide, may not always be needed for an accurate diagnosis.

Whereas most previous research focused on single-test evaluations, the aim of our study was to determine the added diagnostic value of indicators from patient and symptoms and urinalysis, taking into account their mutual dependencies and information from prior tests. Using our results, we present an efficient, easy-to-use diagnostic rule, consisting of a limited number of tests while preserving diagnostic accuracy.

METHODS

Participants

We recruited patients from April 18, 2006, until October 8, 2008, into a cross-sectional study of 20 general practices in and around Amsterdam, The Netherlands. Female patients older than 12 years who contacted their family physician with painful and/or frequent micturition were eligible. Their symptoms had to be present for no longer than 7 days. Exclusion criteria were pregnancy, lactation, signs of pyelonephritis, having used antibiotics or having undergone a urological procedure in the past 2 weeks, known anatomic or functional abnormalities of the urogenital tract, and being immunocompromised (with the exception of diabetes mellitus).

Assessments

Included patients completed a detailed questionnaire to record presence and severity of signs and symptoms (history) on a 4-point scale, and they collected a urine sample at the physician visit. In line with the national guideline of the Dutch College of General Practitioners,[3] no instructions for the collection method were given, because the method of collection has been reported to have no effect on the extent of contamination.[25-27]

The family physician or medical assistant performed a urine dipstick test (Multistix 5, Siemens Medical Solutions Diagnostics) using a Clinitek Status analyzer (Siemens Medical Solutions Diagnostics), as well as a dipslide test (Uricult classic, Orion diagnos-

tica) according to the manufacturer's instructions. The results of these 2 tests were recorded on a standardized registration form.

Immediately after the urine dipstick and dipslide investigations had been performed, urine samples were stored in a refrigerator. Within 8 hours, a specialized courier service (Ruwiel Labexpress, Kockengen, The Netherlands) transported the samples at 4°C to the Academic Medical Center in Amsterdam. A urinary sediment investigation was performed at the Laboratory for Clinical Chemistry, and the urine was cultured at the Laboratory for Medical Microbiology by trained laboratory technicians who had no knowledge of previous test results.

All recorded test results were entered into a structured database by a qualified data entry service (Service Point Nederland BV).

Statistical Analysis

We determined that if the 5 events per variable rule was applied to a number of 20 possible diagnostic indicators, we would need about 100 women with a UTI.[28]

Missing values were imputed using multiple imputation by chained equations.[29,30] Logistic regression with bootstrapped backward elimination was used to select a parsimonious set of variables (P remove .05, bootstrap inclusion fraction ≥66.67%; see the Supplemental Appendix, available at http://annfammed.org/content/11/5/442/suppl/DC1 for details). The result of the urine culture was the binary dependent variable, with more than 10^3 colony-forming units (CFUs) of a single uropathogen per milliliter (mL) being defined as a positive culture according to international guidelines.[31] Different indicators from the same diagnostic medium (eg, nitrite and leukocyte esterase from urine dipstick, or bacteria and leukocytes in sediment) were analyzed as separate variables. Because we hypothesized that women who thought they had a UTI might have a higher UTI probability if they had experienced a (proven) UTI in the past, we entered 2 interaction variables into the backward elimination procedure: the combination of a woman thinking she had a UTI and the number of UTIs she reported to have experienced in the past year, as well as the combination of a woman thinking she had a UTI and reporting to have had at least 1 UTI ever diagnosed by a physician.

Using the selected variables, that is, those retained after backward elimination, we composed 5 different models: (1) history only, (2) history + dipstick, (3) history + dipstick + sediment, (4) history + dipstick + dipslide, and (5) history + dipstick + sediment + dipslide.

We performed parameter-wise shrinkage of the obtained regression coefficients to correct for possible overoptimism.[32] Using the shrunk regression coeffi-

cients, predicted risks were calculated for the 5 different models. Based on 2 independent polls of 150 Dutch family physicians, we considered predicted risks below 30% and above 70% to be clinically relevant. These 2 risk thresholds were used to compare the diagnostic performances of the different models. In addition, we repeated the analysis for risk thresholds of 20% and 80%.

We also composed a sixth model in which a diagnostic step (eg, history, dipstick, sediment, and dipslide) was performed only for those patients who remained in the intermediate risk category (between 30% and 70%) after the preceding step.

For the most informative models, the shrunk regression coefficients were used to compose clinical scores.[33]

Because urinalysis results might be influenced by bladder incubation time,[34] we asked participating patients whether they had been urinating within 4 hours before urine collection and explored whether the diagnostic performance of our model was affected by correcting for this possibility.

Analyses were performed in Stata/SE 10.1 (StataCorp LP).

The study procedure was approved of by the Medical Research Ethics Committee of the Academic Medical Center in Amsterdam. Participating women received a letter with information about the study and provided written informed consent. For patients younger than 18 years, written parental authorization was obtained.

RESULTS

A total number of 205 women were included in the study. Because 92 (45%) had data missing for at least 1 variable, we created 45 imputed data sets.[30,35] No urine culture was available for 9 patients; these patients were dropped from the analysis after multiple imputations,[36] and the resulting 196 patients were used for the analysis.

The patients' main history characteristics are displayed in Table 1. Their

Table 1. Distribution of Patient Characteristics: History

Characteristic	No. (% of Total)	Culture Sample, No.		Positive Culture %
		Positive	Negative	
Total	196 (100)	120	76	61
Ethnicity				
Dutch/German	171 (87)	107	64	63
Surinam/Antillian/Aruban	15 (8)	6	9	40
Other	10 (5)	7	3	70
General health				
Very good	22 (11)	12	10	55
Good	110 (56)	74	36	67
Reasonable	44 (22)	25	19	57
Moderate	11 (6)	7	4	64
Bad	9 (5)	2	7	22
Marital status				
Married	81 (41)	51	30	63
Partner, cohabitating	31 (16)	20	11	65
Partner, living apart	36 (18)	24	12	67
Single	48 (24)	25	23	52
≥ First-grade relative with recurrent UTIs?				
No	137 (70)	83	54	61
Yes	59 (30)	37	22	63
Last menstruation >1 year ago?				
No	133 (68)	81	52	61
Yes	63 (32)	39	24	62
Diabetes mellitus according to patient?				
No	172 (88)	111	61	65
Yes	24 (12)	9	15	38
Times of sexual activity in past week				
0	67 (34)	34	33	51
1	39 (20)	23	16	59
2	40 (21)	33	7	83
3	22 (11)	14	8	64
4	14 (7)	9	5	64
≥5	14 (7)	7	7	50
UTIs in past year according to patient				
0	92 (47)	57	35	62
1	28 (14)	24	4	86
2	30 (15)	19	11	63
≥3	32 (16)	15	17	47
Don't know	14 (7)	5	9	36
≥1 UTI ever diagnosed according to patient?				
No	42 (21)	24	18	57
Yes	154 (79)	100	54	65
Patient thinks she has a UTI?				
No	9 (5)	4	5	44
Yes	164 (84)	114	50	70
Don't know	23 (12)	2	21	9

UTI = urinary tract infection.

Table 1. Distribution of Patient Characteristics: History (continued)

Characteristic	No. (% of Total)	Culture Sample, No.		Positive Culture %[a]
		Positive	Negative	
Symptoms				
Duration of symptoms, d				
0	18 (9)	12	6	67
1	40 (20)	28	12	70
2	37 (19)	23	14	62
3	37 (19)	22	15	59
4	24 (12)	14	10	58
≥5	40 (20)	21	19	53
False urge to urinate				
No	54 (28)	30	24	57
A little	82 (42)	53	29	65
Quite much	43 (22)	27	16	63
Very much	17 (9)	10	7	59
More frequent micturition than usual				
No	12 (6)	3	9	25
A little	62 (32)	31	31	50
Quite much	81 (42)	59	22	73
Very much	41 (21)	27	14	66
Pain during micturition				
No	35 (18)	14	22	39
A little	73 (37)	39	34	53
Quite much	64 (33)	53	11	83
Very much	23 (12)	14	9	61
Urge to urinate hard to control				
No	49 (25)	22	27	45
A little	88 (45)	57	31	65
Quite much	33 (17)	24	9	73
Very much	26 (13)	17	9	65
Vaginal discharge				
No	133 (68)	86	47	65
A little	49 (25)	27	22	55
Quite much	14 (7)	7	7	50
Very much	0 (0)	0	0	n/a
Vaginal irritation or itching				
No	110 (56)	72	38	65
A little	57 (29)	28	29	49
Quite much	19 (10)	15	4	79
Very much	10 (5)	5	5	50
Bother at work/school				
No	24 (13)	9	15	38
Hardly	37 (19)	21	16	57
Moderate	69 (35)	48	21	70
Much	46 (23)	29	17	63
Very much	20 (10)	13	7	65
Bother at social activities				
No	47 (24)	24	23	51
Hardly	49 (25)	29	20	59
Moderate	53 (27)	36	17	68
Much	33 (17)	21	12	64
Very much	14 (7)	10	4	71

n/a = not applicable; UTI = urinary tract infection.

[a] Percentages represent the fraction of positive cultures for the corresponding characteristic, eg, 70% of patients with symptoms for 1 day had a positive culture (implying that the positive predictive value of having symptoms for 1 day is 70%).

mean age was 43 years (range = 16 to 89 years); the mean age did not differ between patients with and those without a positive culture. The characteristics of their urinalyses are displayed in Table 2. The prevalence of a positive culture ($\geq 10^3$ CFU/mL of a uropathogen) was 61% (120/196). Of the 120 positive cultures, 115 had a single uropathogen, and 5 had mixed flora, all of which contained $\geq 10^3$ CFU/mL of the primary uropathogen *Escherichia coli*. Of the 115 single uropathogens, 4 were secondary pathogens (1 had between 10^3 and 10^4 CFU/mL and 3 $\geq 10^4$ CFU/mL).

Seven variables were retained after bootstrapped backward elimination: 3 history variables (having at least considerable pain during micturition, having any vaginal irritation, suspecting a UTI); 2 dipstick variables (nitrite positive, blood ≥1+); 1 sediment variable (>20 leukocytes per high-power field [HPF]); and 1 dipslide variable (cystine lactose electrolyte deficient [CLED] medium $\geq 10^5$ CFU/mL). These 7 variables were used to compose the 6 predefined models.

The 2 interaction variables—the combination of suspecting a UTI and reported number of UTIs in the past year and the combination of suspecting a UTI and reporting at least 1 UTI ever diagnosed by a physician—were not retained (bootstrap inclusion fractions 37.49% and 21.01%, respectively).

Figure 1 shows the discriminative performances of models that were based on analyses for all 196 included patients. After applying the history-only model, 28 patients (14%) had a less than 30% predicted risk and 81 patients (41%) had a greater than 70% predicted risk of UTI. Observed risks were 0.14 and 0.81, respectively. After applying the history and dipstick model, 68 patients (35%) had a less than 30% predicted risk, and 76 patients (39%) had a greater than 70% predicted risk of UTI. Observed risks were 0.21 and 0.95, respectively. Of the 81 patients who had a greater than 70% predicted risk after applying

the history-only model, 35 were classified below that threshold by the history + dipstick model because of a negative nitrite test result, 24 of whom had a positive culture (observed risk 0.68). These 24 patients were correctly classified with the history + dipstick + sediment model, the history + dipstick + dipslide model, or the history + dipstick + sediment + dipslide model.

Figure 2 shows the discriminative performances of the model in which a diagnostic step (eg, history + dipstick + sediment or dipslide) was performed only for those patients who remained in the intermediate risk category (between 30% and 70%) after the preceding step. After performance of history and a urine dipstick,

62 patients (32%) had a predicted risk of less than 30%, 100 patients (51%) had a predicted risk of greater than 70%, and 34 patients (17%) remained in the intermediate risk category because of the combination of a negative nitrite test and a positive blood test. Subsequent performance of either a sediment or a dipslide test reclassified 5% or 7% of patients, respectively, from the intermediate category into the high-risk category (greater than 70%) because of a positive test result. All predicted risks were close to the observed risks, indicating good calibration of all models. Models that included history plus dipstick, sediment, and dipslide, individually and in combination, however, calibrated better than the model that additionally tested patients with intermediate risk, implying a lower rate of false positives and false negatives in the highest and lowest risk category, respectively.

Because we considered history-only or history + dipstick models, the clinically best applicable models for all patients, we used their regression coefficients to compose clinical scores. Table 3 displays the odds ratios, coefficients, and clinical scores for the variables in these 2 models, and Table 4 displays their predicted risks within each predefined risk category (less than 30%, 30% to 70%, and greater than 70%). Reporting any vaginal irritation reduced the probability of a UTI, whereas presence of the other 6 history and dipstick indicators increased this probability. Suspecting a UTI and a positive nitrite test were the strongest indicators of a positive culture. In contrast with history only, having at least considerable pain during micturition did not add any value to the history + dipstick model.

We repeated the analysis for predicted risk thresholds of 20% and 80%, which assigned 31% of patients (60 of 196) to either the high- or the low-risk category after history only (compared with 56% (109 of 196) for thresholds of 30% and 70%), without yielding better calibration (proportion of women with UTI in the high-risk category and women without UTI in the low-risk category).

Table 2. Univariate Distributions of Patient Characteristics: Urinalysis

Characteristic	No. (% of Total)	Culture Result, No.		Positive Culture, %[a]
		Positive	Negative	
Total	196 (100)	120	76	61
Dipstick				
Blood				
Negative	35 (18)	11	24	31
Trace	38 (20)	15	23	39
1+	29 (14)	20	9	69
2+	36 (19)	26	10	72
3+	58 (30)	48	10	82
Leukocytes				
Negative	39 (20)	10	29	26
Trace	15 (8)	5	10	33
1+	47 (24)	31	16	66
2+	35 (18)	27	8	77
3+	60 (31)	47	13	78
Nitrate positive				
No	133 (68)	60	73	45
Yes	62 (32)	60	3	95
Sediment				
Bacteria, No./HPF				
None	52 (27)	18	34	35
Few	48 (24)	27	17	56
Many	53 (27)	40	13	75
Very many	43 (22)	35	8	81
Leukocytes, No./HPF				
0	26 (13)	6	20	23
1-3	32 (16)	9	23	28
4-10	18 (9)	9	9	50
11-20	24 (12)	14	10	58
≥20	96 (49)	82	14	85
Dipslide				
CLED medium, CFU/mL				
$<10^3$	64 (33)	21	43	33
10^3-10^4	32 (16)	10	22	31
10^4-10^5	10 (5)	6	4	60
$≥10^5$	90 (46)	83	7	92

CFU = colony-forming unit; CLED = cystine lactose electrolyte deficient; HPF = high-power field.

[a] Percentages represent the fraction of positive cultures for the corresponding characteristic.

Figure 1. Discrimination by subsequent tests when performed for all patients.
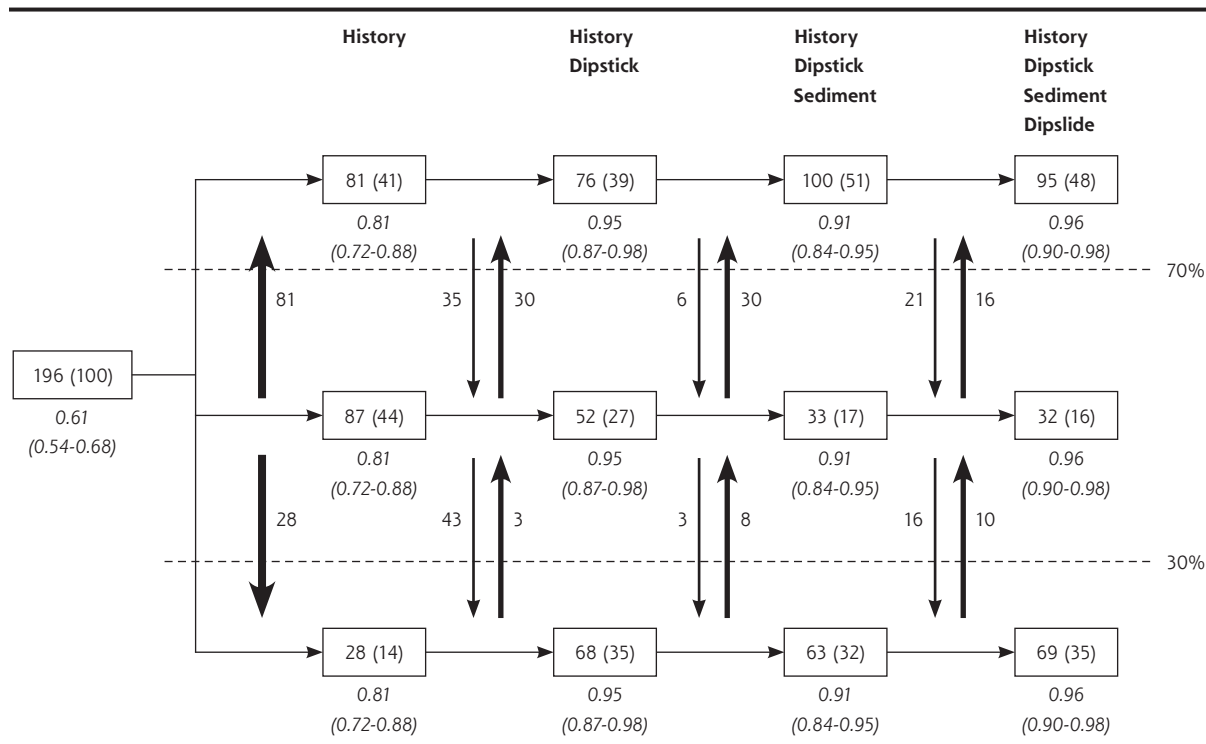
UTI = urinary tract infection.

Notes: In the boxes, patient numbers in each predicted risk category (with corresponding percentage in brackets) are shown after application of each subsequent diagnostic model to all 196 patients. The predicted risk categories used are <30%, 30%-70% and >70%. The vertical arrows display the numbers of patients that switch between categories after performing a test. The observed risks (and their 95% CIs) are shown in italics.

The figure is best read from left to right: 196 patients (100%) were included and the prevalence of UTI was 61%. Based on the results from the history questions only, 28 patients (14%) had a predicted risk of <30% and 81 patients (41%) had a predicted risk of >70%. When both history and a urine dipstick were performed, 68 patients (35%) had a predicted risk of <30% and 76 patients (39%) had a predicted risk of >70%. Additional performance of a sediment and dipslide resulted in removal from the intermediate category of 20 more patients (11%), mainly those who were correctly classified into the highest risk category after history only, but were incorrectly reclassified into the intermediate risk category after a negative nitrite test result on a urine dipstick. All observed risks were close to the predicted risks (good calibration).

Bladder incubation time was 4 or more hours in 26 patients, of whom 20 had a positive culture. Omitting these patients from the overall analysis did not affect the performance of our models.

DISCUSSION

Asking 3 questions (Does the patient think she has a UTI? Is there at least considerable pain on micturition? Is there vaginal irritation?) may be sufficient to correctly classify more than one-half of women with painful and/or frequent micturition as having a UTI risk of either less than 30% or greater than 70%. Subsequent performance of 2 urine dipstick tests (nitrite and blood) raises this proportion to 73% (Figure 1). This percentage rises to 83% if a urine dipstick is performed only for patients with a UTI risk between 30% and 70% after history (Figure 2) and avoids the possibility of a false-negative nitrite test in patients with a high UTI risk (greater than 70%) after history. In our
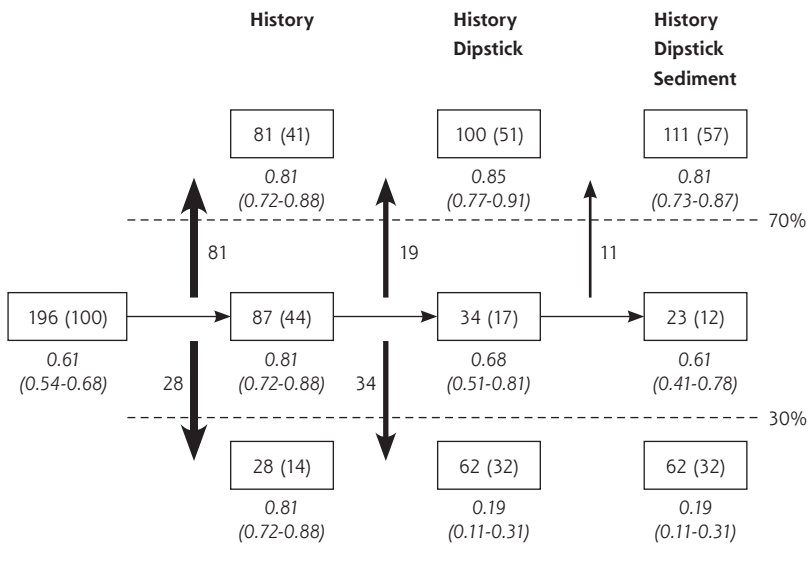
sample, the proportion of false-negative dipstick results among women with a high UTI risk after history was 69% (24/35).

More than 10% of patients will not be classified as having a probability of UTI of less than 30% or greater than 70%, even if all available tests are performed (Figure 1), mainly because of a combined negative nitrite test and a positive blood test on urine dipstick investigation. These women may require a urine culture and, in the case of a negative culture and positive blood test, follow-up of their hematuria.[37]
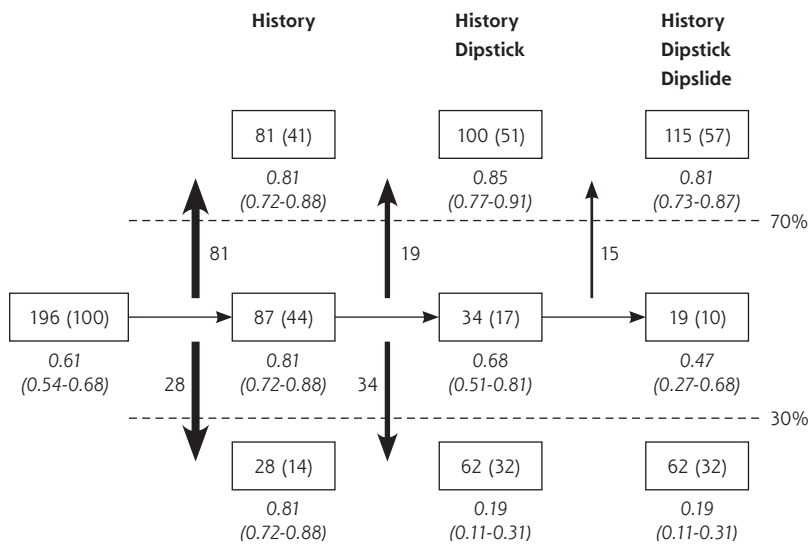
Suspecting a UTI and a positive nitrite test are the strongest indicators of a positive culture. We found that a combination of suspecting a UTI and reported number of UTIs in the past year or a combination of suspecting a UTI and reporting at least 1 UTI ever diagnosed by a physician did not contain any additional information compared with the cumulative yield of the separate indicators (that is, no statistical interaction). In other words, the likelihood of women who

Figure 2. Discrimination by subsequent tests when performed in patients in the intermediate risk category only.

A. Scenario: history + dipstick + sediment.

	History	History Dipstick	History Dipstick Sediment
	81 (41) *0.81 (0.72-0.88)*	100 (51) *0.85 (0.77-0.91)*	111 (57) *0.81 (0.73-0.87)*

70%

↑81 ↑19 ↑11

196 (100) *0.61 (0.54-0.68)* → 87 (44) *0.81 (0.72-0.88)* → 34 (17) *0.68 (0.51-0.81)* → 23 (12) *0.61 (0.41-0.78)*

↓28 ↕34

30%

| 28 (14) *0.81 (0.72-0.88)* | 62 (32) *0.19 (0.11-0.31)* | 62 (32) *0.19 (0.11-0.31)* |

B. Scenario: History + dipstick + dipslide.

	History	History Dipstick	History Dipstick Dipslide
	81 (41) *0.81 (0.72-0.88)*	100 (51) *0.85 (0.77-0.91)*	115 (57) *0.81 (0.73-0.87)*

70%

↑81 ↑19 ↑15

196 (100) *0.61 (0.54-0.68)* → 87 (44) *0.81 (0.72-0.88)* → 34 (17) *0.68 (0.51-0.81)* → 19 (10) *0.47 (0.27-0.68)*

↓28 ↕34

30%

| 28 (14) *0.81 (0.72-0.88)* | 62 (32) *0.19 (0.11-0.31)* | 62 (32) *0.19 (0.11-0.31)* |

UTI = urinary tract infection.

Notes: In the boxes, patient numbers (percent) in each predicted risk category are shown after application of each subsequent diagnostic model to patients in the intermediate predicted risk category only for the scenarios history + dipstick + sediment and history + dipstick + dipslide (Figure 2A and 2B, respectively). The predicted risk categories used are <30%, 30%-70% and >70%. The vertical arrows display the numbers of patients that moved from the intermediate risk category to one of the extreme risk categories (<30% or >70%) after performing a test. The observed risks (95% CI) are in italics.

Figures are best read from left to right: 196 patients (100%) were included, and the prevalence of UTI was 61%. Based on the results from the history questions only, 28 patients (14%) had a predicted risk of <30% and 81 patients (41%) had a predicted risk of >70%. After subsequent performance of a urine dipstick for patients in the intermediate risk category, 62 patients (32%) had a predicted risk of <30%, 100 patients (51%) had a predicted risk of >70%, and 34 patients (17%) remained in the intermediate risk category (between 30% and 70%) as the result of a negative nitrite test and a positive blood test. Subsequent performance of either a sediment or a dipslide reclassified 5% or 7%, respectively, of patients from the intermediate category into the high-risk category (>70%) because of a positive test result (Figure 2A and 2B, respectively). All observed risks were close to the predicted risks (good calibration).

think they have a UTI does not depend on their previous experiences with UTIs.

The diagnostic performances of UTI indicators were previously described.[16,38,39] In contrast with Bent et al, who did not take into account the mutual dependencies of the diagnostic indicators[16] and therefore may have overestimated their predictive values, Little et al and McIsaac et al described multivariable analyses similar to ours.[38,39] Nevertheless, there are some methodological differences between these studies and our study. First, they included women in whom family physicians suspected a UTI, based on their personal judgment, whereas we used clearly formulated eligibility criteria to better ensure generalizability. Second, they transported urine samples before cultures were made, whereas we refrigerated our samples until cultures were made to assure reliability of the reference standard. Third, they did not report how missing values were handled, whereas we gave an extensive description of our multiple imputation method. Fourth, as did Bent et al, they did not assess diagnostic values of sediment and dipslide investigations, whereas we considered evaluation of these investigations an essential part of our analysis and showed their limited added value empirically. Finally, we investigated the scenario of performing a test only for women with an intermediate UTI risk after the preceding test (that is, women with a predicted risk between 30% and 70%). Although we consider this scenario of substantial clinical relevance, we did not present it as our main result because of small patient numbers and the resulting risk of unstable estimates.

To avoid overfitting by using too many candidate predictors, we could not include all available

Table 3. Clinical Scores for History Only and History + Dipstick: Regression Coefficients and Scores of Selected Indicators

Indicator[a]	History Only			History + Dipstick		
	Odds Ratio (95% CI)	Coefficient	Score	Odds Ratio (95% CI)	Coefficient	Score
History						
Patient thinks she has a UTI	8.85 (3.35-24.05)	2.18	8	15.64 (3.94-63.43)	2.75	11
At least considerable pain on micturition	2.80 (1.52-5.16)	1.03	4	0.97 (0.91-1.03)	−0.03	0[b]
Vaginal irritation[c]	0.77 (0.54-1.08)	−0.26	−1	0.61 (0.36-1.04)	−0.50	−2
Dipstick						
Nitrate positive	–	–	–	31.19 (7.39-130.32)	3.44	14
Blood ≥1+	–	–	–	7.32 (3.03-17.81)	1.99	8
Regression intercept	–	−1.77	–	–	−3.89	–

UTI = urinary tract infection.

[a] Indicators from history and urine dipstick that were retained after logistic regression analysis with bootstrapped backward elimination (*P* remove .05).

[b] This clinical score of 0 had no added value.

[c] Vaginal irritation reduced the probability of a positive culture, whereas the other indicators increased it.

Table 4. Clinical Scores for History Only, and History + Dipstick: Sum Scores and Predicted Probabilities per Risk Category

Risk Category	History Only		History + Dipstick	
	Score	Predicted Probability (95% CI)	Score	Predicted Probability (95% CI)
<30%	≤3	16 (7-34)	≤12	15 (7-31)
30%-70%	4-8	56 (44-68)	14-17	61 (41-77)
>70%	≥11	79 (69-86)	≥19	91 (76-96)

Notes: For example, for history only, patients with a sum score of ≥3 have a predicted probability of 16% and are therefore classified into the lowest risk category (<30%).

variables in our analysis; we therefore selected the 22 variables we considered the most relevant based on literature and clinical usefulness (Supplemental Appendix). This number was in accordance with the 5 events per variable rule,[28] as we included 120 patients with a positive culture.

Limitations

Like in any prediction model, our results may be overoptimistic for the population in which the model was developed. As a result, it may perform less well in a different population. Although we restricted the number of candidate indicators to 22 and performed bootstrapped selection and parameter-wise shrinkage, external validation of the model in a different population is indicated.

The diagnostic indicator of having at least considerable pain during micturition might have a cultural dimension that is due to differences in pain experi-

ences between different subcultures. This possibility should be considered when our model is applied to individual patients in clinical practice.

To compare the diagnostic performances of the different models in a way that is attractive to clinicians, we used classification into predicted risk categories instead of traditional performance measures, such as receiver operating characteristic areas, which may be hard to apply in daily practice.[40] Because there are no widely accepted methods to define these risk categories, we performed 2 independent polls of 150 Dutch family physicians and chose to use predicted risk thresholds of less than 30% and greater than 70% based on the results of those polls. Nevertheless, we repeated the analysis for predicted risks of less than 20% and greater than 80%. Although using these more extreme risk thresholds may seem more accurate to either detect or rule out UTI, it did not change the observed risks in either the high- or the low-risk category. Moreover, it resulted in worse discrimination: whereas application of the 30% and 70% thresholds assigned 56% of patients to either the high- or the low-risk category after history only, application of the 20% and 80% thresholds did so for only 31%. Because of the better discriminative properties and the results of the 2 polls, we decided to present the results of the 30%-70% instead of the 20%-80% thresholds.

An uncommon but severe complication of UTIs is pyelonephritis, which may be a reason to treat women at a low probability of UTI. Placebo arms of randomized controlled trials suggest, however, that cystitis seldom progresses to pyelonephritis.[41-43] No patients consistently developed pyelonephritis in our own study population during the week after diagnosis (95% CI, 0%-2%).

An important issue in UTI diagnosis is the cutoff value to be used for defining a culture as being positive. In practice, many clinicians and medical microbiologists still use the traditional cutoff value of 10^5 CFU/mL as described by Kass in a study on asymptomatic women and women with pyelonephritis.[44] For symptomatic women, however, the most recent international guidelines recommend a cutoff value of 10^3 CFU/mL.[31] Because we think that a reference standard in diagnostic research should be based on scientific evidence rather than on practical arguments, we chose to use a cutoff value of 10^3 CFU/mL as the dependent variable in our analysis; however, we performed a separate analysis using a cutoff value of 10^5 CFU/mL value as the dependent variable, which yielded the same diagnostic indicators and values, with the exception of vaginal irritation, which was not selected.

According to the previously mentioned guideline,[31] the cutoff value of $\geq 10^3$ CFU/mL applies only to primary pathogens. For secondary and doubtful pathogens, different cutoff values are recommended ($\geq 10^4$ and $\geq 10^5$ CFU/mL, respectively). We found only 4 secondary pathogens (of which 1 was between $\geq 10^3$ and $\geq 10^4$ CFU/mL, and 3 of which were $\geq 10^4$ CFU/mL) and no doubtful pathogens. As a result, choosing a different cutoff value for these pathogens would not affect our findings.

A factor that might influence urinalysis results is bladder incubation time, which should be preferably 4 or more hours.[34] Because we did not consider incubation time to be a potential diagnostic indicator and because we chose to stay close to common practice (in which bladder incubation is generally hard to determine), we did not include it as a variable in our analysis. We did explore, however, whether the diagnostic performance of our model was affected by correcting for bladder incubation time, which was not the case.

Our findings imply that UTI diagnosis may be simplified by considerably reducing the number of questions and urine investigations needed. In more than one-half of women complaining of painful and/or frequent micturition, the diagnostic procedure may be limited to asking 3 simple questions (Does the patient think she has a UTI? Is there at least considerable pain on Micturition? Is there vaginal irritation?), meaning that it might be completed by telephone in these cases. Additional performance of a nitrite and blood dipstick seems sufficient to make an accurate diagnosis for most patients. Urinary sediment and dipslide appear to add little information to what is already known from history and dipstick results, implying that performance of these expensive, time-consuming tests might be abandoned.

To read or post commentaries in response to this article, see it online at http://www.annfammed.org/content/11/5/442.

References

1. Foxman B. Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. *Dis Mon.* 2003;49(2):53-70.

2. Gonzalez CM, Schaeffer AJ. Treatment of urinary tract infection: what's old, what's new, and what works. *World J Urol.* 1999;17(6):372-382.

3. van Haaren KAM, Visser HS, van Vliet S, et al. NHG-Standaard Urineweginfecties (tweede herziening). *Huisarts Wet.* 2005;48(6):341-352.

4. Warren JW, Abrutyn E, Hebel JR, Johnson JR, Schaeffer AJ, Stamm WE; Infectious Diseases Society of America (IDSA). Guidelines for antimicrobial treatment of uncomplicated acute bacterial cystitis and acute pyelonephritis in women. *Clin Infect Dis.* 1999;29(4):745-758.

5. Foxman B, Barlow R, D'Arcy H, Gillespie B, Sobel JD. Urinary tract infection: self-reported incidence and associated costs. *Ann Epidemiol.* 2000;10(8):509-515.

6. Clayson D, Wild D, Doll H, Keating K, Gondek K. Validation of a patient-administered questionnaire to measure the severity and bothersomeness of lower urinary tract symptoms in uncomplicated urinary tract infection (UTI): the UTI Symptom Assessment questionnaire. *BJU Int.* 2005;96(3):350-359.

7. Colgan R, Keating K, Dougouih M. Survey of symptom burden in women with uncomplicated urinary tract infections. *Clin Drug Investig.* 2004;24(1):55-60.

8. Ellis AK, Verma S. Quality of life in women with urinary tract infections: is benign disease a misnomer? *J Am Board Fam Pract.* 2000;13(6):392-397.

9. Malterud K, Baerheim A. Peeing barbed wire. Symptom experiences in women with lower urinary tract infection. *Scand J Prim Health Care.* 1999;17(1):49-53.

10. Barry HC, Ebell MH, Hickner J. Evaluation of suspected urinary tract infection in ambulatory women: a cost-utility analysis of office-based strategies. *J Fam Pract.* 1997;44(1):49-60.

11. Fenwick EA, Briggs AH, Hawke CI. Management of urinary tract infection in general practice: a cost-effectiveness analysis. *Br J Gen Pract.* 2000;50(457):635-639.

12. Chomarat M. Resistance of bacteria in urinary tract infections. *Int J Antimicrob Agents*. 2000;16(4):483-487.

13. Gupta K, Hooton TM, Stamm WE. Increasing antimicrobial resistance and the management of uncomplicated community-acquired urinary tract infections. *Ann Intern Med*. 2001;135(1):41-50.

14. Hooton TM, Besser R, Foxman B, Fritsche TR, Nicolle LE. Acute uncomplicated cystitis in an era of increasing antibiotic resistance: a proposed approach to empirical therapy. *Clin Infect Dis*. 2004; 39(1):75-80.

15. Kahlmeter G; ECO.SENS. An international survey of the antimicrobial susceptibility of pathogens from uncomplicated urinary tract infections: the ECO.SENS Project. *J Antimicrob Chemother*. 2003;51 (1):69-76.

16. Bent S, Nallamothu BK, Simel DL, Fihn SD, Saint S. Does this woman have an acute uncomplicated urinary tract infection? *JAMA*. 2002;287(20):2701-2710.

17. Devillé WL, Yzermans JC, van Duijn NP, Bezemer PD, van der Windt DA, Bouter LM. The urine dipstick test useful to rule out infections. A meta-analysis of the accuracy. *BMC Urol*. 2004;4:4.

18. Medina-Bombardó D, Seguí-Díaz M, Roca-Fusalba C, Llobera J; dysuria team. What is the predictive value of urinary symptoms for diagnosing urinary tract infection in women? *Fam Pract*. 2003;20(2):103-107.

19. Winkens R, Nelissen-Arets H, Stobberingh E. Validity of the urine dipslide under daily practice conditions. *Fam Pract*. 2003;20(4): 410-412.

20. Begg CB. Biases in the assessment of diagnostic tests. *Stat Med*. 1987; 6(4):411-423.

21. Dawid AP. Properties of diagnostic data distributions. *Biometrics*. 1976;32(3):647-658.

22. Feinstein AR. Clinical biostatistics. XXXIX. The haze of Bayes, the aerial palaces of decision analysis, and the computerized Ouija board. *Clin Pharmacol Ther*. 1977;21(4):482-496.

23. Lachs MS, Nachamkin I, Edelstein PH, Goldman J, Feinstein AR, Schwartz JS. Spectrum bias in the evaluation of diagnostic tests: lessons from the rapid dipstick test for urinary tract infection. *Ann Intern Med*. 1992;117(2):135-140.

24. Lofgren RP. The dynamic nature of sensitivity and specificity. *J Gen Intern Med*. 1987;2(6):452-453.

25. Baerheim A, Laerum E. Home-voided urine specimens in women. Diagnostic agreement with clean-catch midstream specimens. *Scand J Prim Health Care*. 1990;8(4):207-211.

26. Leisure MK, Dudley SM, Donowitz LG. Does a clean-catch urine sample reduce bacterial contamination? *N Engl J Med*. 1993;328(4): 289-290.

27. Lifshitz E, Kramer L. Outpatient urine culture: does collection technique matter? *Arch Intern Med*. 2000;160(16):2537-2540.

28. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. *Am J Epidemiol*. 2007;165(6): 710-718.

29. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res*. 2007;16(3): 219-242.

30. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med*. 2011;30(4): 377-399.

31. European Confederation of Laboratory Medicine. European urinalysis guidelines. *Scand J Clin Lab Invest Suppl*. 2000;231:1-86.

32. Verweij PJ, Van Houwelingen HC. Cross-validation in survival analysis. *Stat Med*. 1993;12(24):2305-2314.

33. Sullivan LM, Massaro JM, D'Agostino RB Sr. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Stat Med*. 2004;23(10):1631-1660.

34. Wilson ML, Gaido L. Laboratory diagnosis of urinary tract infections in adult patients. *Clin Infect Dis*. 2004;38(8):1150-1158.

35. Bodner TE. What Improves with Increased Missing Data Imputations? *Struct Equ Model: Multidiscip J*. 2008;15(4):651-675.

36. von Hippel PT. Regression with missing Ys: An improved strategy for analyzing multiply imputed data. *Sociol Methodol*. 2007; 37(1):83-117.

37. Kelly JD, Fawcett DP, Goldberg LC. Assessment and management of non-visible haematuria in primary care. *BMJ*. 2009;338:a3021.

38. Little P, Turner S, Rumsby K, et al. Developing clinical rules to predict urinary tract infection in primary care settings: sensitivity and specificity of near patient tests (dipsticks) and clinical scores. *Br J Gen Pract*. 2006;56(529):606-612.

39. McIsaac WJ, Low DE, Biringer A, Pimlott N, Evans M, Glazier R. The impact of empirical management of acute cystitis on unnecessary antibiotic use. *Arch Intern Med*. 2002;162(5):600-605.

40. Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. *Clin Chem*. 2008;54(1):17-23.

41. Christiaens TC, De Meyere M, Verschraegen G, Peersman W, Heytens S, De Maeseneer JM. Randomised controlled trial of nitrofurantoin versus placebo in the treatment of uncomplicated urinary tract infection in adult women. *Br J Gen Pract*. 2002;52(482):729-734.

42. Ferry SA, Holm SE, Stenlund H, Lundholm R, Monsen TJ. Clinical and bacteriological outcome of different doses and duration of pivmecillinam compared with placebo therapy of uncomplicated lower urinary tract infection in women: the LUTIW project. *Scand J Prim Health Care*. 2007;25(1):49-57.

43. Richards D, Toop L, Chambers S, Fletcher L. Response to antibiotics of women with symptoms of urinary tract infection but negative dipstick urine test results: double blind randomised controlled trial. *BMJ*. 2005;331(7509):143.

44. Kass EH. Bacteriuria and the diagnosis of infections of the urinary tract; with observations on the use of methionine as a urinary antiseptic. *AMA Arch Intern Med*. 1957;100(5):709-714.